UNIVERSITI TUNKU ABDUL RAHMAN
Department of Mathematics and Actuarial Science

CONTENTS

1 Introduction to the Related Statistical Concepts

1.1 Conditional Probability Distribution

Definition 1.
If X and Y are discrete random variables with joint probability function $p(x, y)$ and marginal probability functions $p_1(x)$ and $p_2(y)$, respectively, then the **conditional discrete probability function** of X given Y is

$$\begin{aligned} p(x|y) &= P(X = x|Y = y) \\ &= \frac{P(X=x,Y=y)}{P(Y=y)} \\ &= \frac{p(x,y)}{p_2(y)} \end{aligned}$$

provided that $p_2(y) > 0$.

$$\boxed{p(x|y) = \frac{p(x,y)}{p_2(y)}}$$

Definition 2.
Let X and Y be jointly continuous random variables with joint density function $f(x, y)$ and marginal densities functions $f_1(x)$ and $f_2(y)$, respectively. For any y such that $f_2(y) > 0$, the conditional density of X given $Y = y$ is given by

$$\boxed{f(x|y) = \frac{f(x,y)}{f_2(y)}}$$

and, for any x such that $f_1(x) > 0$, the conditional density of Y given $X = x$ is given by

$$\boxed{f(y|x) = \frac{f(x,y)}{f_1(x)}}$$

Example 1.
Suppose that $p(x, y)$, the joint probability function of X and Y, is given by

$p(30, 10) = 0.36$, $p(30, 20) = 0.24$,
$p(40, 10) = 0.12$, $p(40, 20) = 0.28$
Calculate the probability function of X given that $Y = 20$.

Example 2.
Let X and Y have joint probability density function(pdf) $f(x, y) = ce^{-4y}, 0 < x < y < \infty$ and zero otherwise. Find the conditional pdf of $Y|X = 140$.

Example 3.
Suppose the joint density of Θ and X is given by

$$f(x,\theta) = \begin{cases} 9\theta^{-5}xe^{-(3+x)/\theta}, & x > 0, \theta > 0 \\ 0 & \text{otherwise} \end{cases}$$

Derive the conditional density of $X = x$ given $\Theta = \theta$. $\boxed{\frac{xe^{-\frac{x}{\theta}}}{\theta^2}}$

Example 4.
Suppose the joint density function of X and Θ is given by

$$f(x,\theta) = \begin{cases} 20\theta^{-4}(1-\theta)^2x(\theta-x)^3, & 0 < x < \theta, < 1 \\ 0, & \text{otherwise} \end{cases}$$

Derive the conditional density of X given $\Theta = \theta$. $\boxed{20\theta^{-5}x(\theta-x)^3}$

1.2 Conditional Expectations

Definition 3.
If X and Y are any two random variables, the conditional expectation of $g(X)$, given that $Y = y$, is define to be

$$\boxed{E(g(X)|Y = y) = \int_{-\infty}^{\infty} g(x)f(x|y)dx}$$

if X and Y are jointly continuous and

$$\boxed{E(g(X)|Y = y) = \sum_{-\infty}^{\infty} g(x)p(x|y)}$$

if X and Y are jointly discrete.

Example 5.
Suppose that $p(x, y)$, the joint probability function of X and Y, is given by

$p(30, 10) = 0.36$, $p(30, 20) = 0.24$,
$p(40, 10) = 0.12$, $p(40, 20) = 0.28$

(i) Find $E(X|Y = 20)$ $\boxed{35.3846}$

(ii) $E(e^X|Y = 20)$. $\boxed{1.2675 \times 10^{17}}$

Example 6 (T1Q1).

Let X and Y have joint probability density function(pdf) $f(x, y) = ce^{-3y}, 0 < x < y < \infty$ and zero otherwise. Find the mean of the conditional distribution of $Y|X = 2.55$.

Example 7 (T1Q2).

Let X_1 and X_2 be independent random variables. X_1 follows a gamma distribution with parameters $\alpha = 3$ and $\beta = \frac{1}{10}$, whereas X_2 follows an exponential distribution with mean $\frac{1}{7}$. Find $E[X_1|X_1 + X_2 = 3]$.

Theorem 1.
Let X and Y denote random variables. Then

$$\boxed{E(X) = E[E(X|Y)]}$$

where, on the right hand side, the inside expectation is with respect to the conditional distribution of X given Y, and the outside expectation is with respect to the distribution of Y.

Theorem 2.
Let X and Y denote random variables and $h(x, y)$ is a function. Then

$$E[h(X,Y)] = E_Y[E(h(X,Y)|Y)]$$

or

$$E[h(X,Y)] = E_X[E(h(X,Y)|X)]$$

Theorem 3.
Let X and Y denote random variables. Then

$$\boxed{V(X) = E[V(X|Y)] + V[E(X|Y)]}$$

Example 8.
Claim size is exponentially distributed with mean λ. λ varies by insured, and follows a Pareto distribution with parameters $\alpha = 5$ and $\theta = 6$. Calculate the expected value of claim size.[1.5]

Example 9. [T1Q3]

You are given the followings:
A portfolio of risks consists of 2 classes, A and B. For an individual risk in either class, the number of claims has the following distribution.

Class	Number of Exposure	Distribution of Claim Frequency	
		Mean	S. Deviation
A	500	0.05	0.248
B	500	0.22	0.625
Total Portfolio	1,000		

Determine the standard deviation of the claim frequency for the total portfolio.

Example 10.

The number of claims on a policy has a Poisson distribution with mean P. P varies by policyholder. P is uniformly distributed on $[9, 11]$. Calculate the variance of the number of claims.

Example 11. (**The Expectation and Variance of a Random Number of Random Variables**)

Suppose $X_1,\ X_2, ...$ are independent and identically distributed, and if N is a nonnegative integer valued random variable independent of $X's$, then find

(a.) $E[\sum_{i=1}^{N} X_i]$ $\boxed{E(N)E(X)}$

(b.) $V[\sum_{i=1}^{N} X_i]$ $\boxed{E(N)V(X)+V(N)E^2(X)}$

Example 12. (**Compound Poisson Distribution**)
Suppose the claim frequency per period N has a Poisson distribution with mean λ, and the claim severity distribution is X, with mean μ_X and variance σ_X^2, and suppose that N and the $X's$ are independent. The aggregate claim per period is $S = X_1 + X_2 + ... + X_N$. Find

(a.) $E(S)$ $\boxed{\lambda\mu_X}$

(b.) $V(S)$ $\boxed{\lambda(\mu_X^2 + \sigma_X^2)}$

1.3 Mixed Distributions

1.3.1 Finite Mixture Distribution

Definition 4.
A random variable Y is a k-point mixture of random variables $X_1, X_2, ..., X_k$ if its cdf is given by

$$\boxed{F_Y(y) = w_1 F_{X_1}(y) + w_2 F_{X_2}(y) + \cdots + w_k F_{X_k}(y)}$$

where all $w_j > 0$ and $w_1 + w_2 + ... + w_k = 1$.

Since the density function is the derivative of the distribution function, it would be the same weighted average of the individual density functions, i.e.,

$$\boxed{f_Y(y) = w_1 f_{X_1}(y) + w_2 f_{X_2}(y) + \cdots + w_k f_{X_k}(y)}$$

Example 13.
Suppose X is a mixture of an exponential distribution with mean 100 and a Pareto distribution with parameters $\alpha = 2, \theta = 200$, and the weight are 60% exponential, 40% Pareto. Find the cdf of X.

1.3.2 Continuous Mixing

Theorem 4.
Let X have pdf $f_{X|\Lambda}(x|\lambda)$ and cdf $F_{X|\Lambda}(x|\lambda)$, where λ is a parameter of X, while X may have other parameters, there are not relevant. Let λ be a realization of the random variable Λ with pdf $\pi_{|\Lambda}(\lambda)$. Then the unconditional pdf of X is

$$\boxed{f_X(x) = \int f_{X|\Lambda}(x|\lambda)\pi_\Lambda(\lambda)d\lambda}$$

where the integral is taken over all values of λ with positive probability.

The resulting distribution is a *mixture distribution*. The distribution function can be determined from

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \int f_{Y|\Lambda}(y|\lambda)\pi_\Lambda(\lambda)d\lambda dy \\ &= \int \int_{-\infty}^{x} f_{Y|\Lambda}(y|\lambda)\pi_\Lambda(\lambda)dyd\lambda \\ &= \boxed{\int F_{X|\Lambda}(x|\lambda)\pi_\Lambda(\lambda)d\lambda} \end{aligned}$$

Similarly, $\boxed{S_X(x) = \int S_{X|\Lambda}(x|\lambda)\pi_\Lambda(\lambda)d\lambda}$

Example 14 (T1Q4)**.**
Given a value of $(\Theta = \theta)$, the random variable X follows a Gamma distribution with probability density function

$$f(x) = \theta^2 x e^{-\theta x}.$$

Θ has a uniform distribution on the interval (2, 12). Determine $S_X(0.25)$ for the unconditional distribution. .

Example 15 (T1Q5).

Annual claim counts per risk are binomial with parameter $m = 3$ and Q. Q varies by risk uniformly on $(0.33, 0.83)$. For a risk selected at random, determine the probability of at most one claims.

1.3.3 Moments of the Mixture Distribution

$$\begin{aligned}
E(Y^m) &= \int y^m f_Y(y)dy \\
&= \int y^m[w_1 f_{X_1}(y) + ... + w_k f_{X_k}(y)]dy \\
&= w_1 \int y^m f_{X_1}(y)dy + ... + w_k \int y^m f_{X_k}(y)dy \\
&= w_1 E(X_1^m) + ... + w_k E(X_k^m) \\
&= \boxed{\sum_{i=1}^k w_i E(X_i^m)}
\end{aligned}$$

for Finite Mixture Distribution

$$\begin{aligned}
E(X^m) &= \int X^m f_X(x)dx \\
&= \int X^m \int f_X(X|\lambda)\pi(\lambda)d\lambda dx \\
&= \int \int X^m f_X(X|\lambda)\pi(\lambda)dx d\lambda \\
&= \int E(X^m|\lambda)d\lambda \\
&= \boxed{E[E(X^m|\lambda)]}
\end{aligned}$$

for Continuous Mixture Distribution

In general,

$$\boxed{E(X^m) = E_\Lambda[E_X(X^m|\lambda)]}$$

and, in particular,

$$\boxed{V(X) = E[V(X|\lambda)] + V[E(X|\lambda)].}$$

Example 16.

On a auto liability coverage, there are three classes of policyholders, young, middle-age and old. 18%;, 62%;, and 20%; of drivers are in young, middle-age and old class. The distribution of the amount of losses for the drivers are:

	Amount of Losses		
Class	500	5,000	10,000
Young	0.09	0.21	0.7
Middle-Age	0.15	0.27	0.58
old	0.59	0.11	0.3

A claim is submitted by a randomly selected driver. Calculate the variance of size of the claim.

Example 17 (T1Q6)**.**

You are given:

- Conditional on λ, the random variables X_1, X_2,. . .,X_m, are independent and follow a Poisson distribution with parameter λ.
- $S_m = X_1 + X_2 + \cdots + X_m$
- The distribution of λ is Gamma with parameters $\alpha = 7$ and $\theta = 16$.

Determine the variance of the marginal distribution of S_{146}.
.

Example 18 (T1Q7).

The size of loss, X, has mean 4λ and variance $8\lambda^2$. Λ has the following density function:

$$f(\lambda) = 5(1,910)^5/(\lambda + 1,910)^6$$

Calculate the variance of the loss.

1.3.4 Gamma Mixture of Poisson- Negative Binomial

The negative binomial can be derived as a gamma mixture of Poisson.
Suppose

$$X|\Lambda \sim POI(\lambda)$$

and

$$\Lambda \sim Gamma(\alpha, \theta)$$

Then

$$X \sim NB(r = \alpha, \beta = \theta)$$

Example 19 (T1Q8).

The annual number of accidents for an individual driver has a Poisson distribution with mean λ. The mean, Λ, of a heterogeneous population of drivers have a gamma distribution with mean 1.04 and variance 0.2704. Calculate the probability that a driver selected at random from the population will have 2 or more accidents in one year.

1.3.5 Exponential Mixture of Inverse Gamma- Pareto

The Pareto distribution can be derived as a exponential mixture of Inverse Gamma.

If

$$X|\beta \sim \text{Exp}(\beta)$$

and

$$\beta \sim \text{Inverse Gamma}(\alpha, \theta)$$

then

$$X \sim \text{Pareto}(\alpha, \theta)$$

Example 20.

You are given the following:

- The amount of an individual claim, Y, follows an exponential distribution function with probability density function
$$f(y|\delta) = \frac{1}{\delta}e^{-y/\delta}, y, \delta > 0.$$
- The mean claim amount, δ, follows an inverse gamma distribution with density function
$$\pi(\delta) = \frac{5^4}{\Gamma(4)}\frac{e^{-5/\delta}}{\delta^5}, \delta > 0.$$

Determine the unconditional density of Y at $y = 5$.

1.4 The Method of Moments

The k^{th} moment of a random variable, taken about the origin, is
$$\boxed{\mu_k' = E(X^k)}$$
The corresponding k^{th} sample moment is the average
$$\boxed{m_k' = \frac{1}{n}\sum_{i=1}^{n} x_i^k}$$

Method of Moments

Choose as estimates those values of the parameters that are solutions of the equations $\mu_k' = m_k'$, for $k = 1, 2, \ldots, t$, where t is the number of parameters to be estimated.

Example 21.

Consider a random sample of size n from a uniform distribution over the interval $(0, \theta)$.

(a) Use the method of moments to estimate the parameter θ, $\tilde{\theta}$.

(b) Show that $\tilde{\theta}$ is an unbiased estimator of θ.

Example 22. [T1Q9]

The random variable X has the density function with parameter β given by

$$f(x;\beta) = \frac{1}{\beta^5} x^4 e^{-\frac{1}{5}(x/\beta)^5}, x > 0, \beta > 0.$$

You are given the following observationof X:

$$4.2, 1.4, 3.4, 6.1, 4.0.$$

Determine the method of moments estimate of β. [Note: $\Gamma(1 + 1/5) = 0.9182$.]

1.5 Maximum Likelihood Estimation

Definition 5.

1. An observation is ***truncated from below*** (also called left truncated) at d if when it is below d it is not recorded but when it is above d it is recorded at its observed value.
2. An observation is ***truncated from above*** (also called right truncated) at u if when it is above u it is not recorded but when it is below u it is recorded at its observed value.
3. An observation is ***censored from below*** (also called left censored) at d if when it is below d it is recorded as being equal to d but when it is above d it is recorded at its observed value.
4. An observation is ***Censored from above*** (also called right censored) at u if when it is above u it is recorded as being equal to u but when it is below u it is recorded at its observed value.

Notes:

1. The most occurrences are left truncation and right censoring.
2. Examples of left truncations:

   (a) Ordinary deductible, d is applied to a policy, and
   (b) Entry age in a survival analysis.
3. Examples of right censoring:
   (a) Policy limit, and
   (b) A person is still alive in a survival analysis when the study ends.

1.5.1 Complete, Individual Data

When there is no truncation, and no censoring and the value of each observation is recorded. Then

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

and

$$l(\theta) = \sum_{i=1}^{n} lnf(x_i|\theta).$$

The notation indicates that it is not necessary for each observation to come from the same distribution.

Example 23.

You are given the following three observations:

$$0.64, 0.85, 0.9$$

You fit a distribution with the following density function to the data:

$$f(x) = (p+1)x^p, 0 < x < 1, p > -1$$

Determine the maximum likelihood estimate of p.

Example 24.
The number of claims per month Y arising on a certain portfolio of insurance policies is to be modelled using a modified geometric distribution with probability density given by

$$f(y) = \frac{\alpha^{y-1}}{(1+\alpha)^y}, y = 1, 2, 3, \ldots$$

Derive the maximum likelihood estimate of α.

Example 25.
You are given:

$$f(x) = \frac{1}{\beta} e^{-(\frac{x-\tau}{\beta})}, x \geq \tau$$

for $\beta > 0$ and $\tau \in \mathbf{R}$. Suppose that $X_1, X_2, \ldots, X_n$ are iid with pdf $f(x|\beta, \tau)$. For any $(\beta, \tau) \in (0, \infty) \times \mathbf{R}$, the probability is 1 that the likelihood has a maximum. Identify the maximum likelihood estimator, and carefully argue that it does indeed maximize the likelihood.

1.5.2 Complete, Grouped Data

When data are complete and grouped, the observations may be summarized as follows. Begin with a set of numbers $c_0 < c_1 < ... < c_k$, where c_0 is the smallest possible observation(often zero) and c_k is the largest observation (often infinity). From the sample, let n_j be the number of observations in the interval (c_{j-1}, c_j). For such data, the likelihood function is

$$L(\theta) = \prod_{j=1}^{k} [F(c_j|\theta) - F(c_{j-1}|\theta)]^{n_j}$$

and its logarithm is

$$l(\theta) = \sum_{j=1}^{k} n_j ln[F(c_j|\theta) - F(c_{j-1}|\theta)]$$

Example 26 (T1Q10).

You are given the following data for claim sizes:

Claim size	Number of claims
Under 1100	11
[1100, 2200)	5
2200 and up	2

The data are fit to an exponential distribution using maximum likelihood. Determine the fitted mean.

1.5.3 Censoring

For censor data from above at u(Right censor), such as data in the presence of claims limit, treat it like grouped data: the likelihood function is the probability of being beyond the censoring point, i.e. $1 - F(u)$.

If the data are left censored, that is, you know an observation is below d but you don't know its exact value, then the likelihood is $F(d)$.

Example 27. [T1Q11]

An auto liability coverage has a claims limit of 110. Claim sizes observed are

$$24, \quad 46, \quad 58, \quad 88, \quad 110$$

where the claim at 110 was for exactly 110. In addition, there are 2 claims above the limit. The data are fitted to an exponential distribution. Determine the MLE of θ.

Example 28 (T1Q12).

Annual claim counts follow a geometric distriution with mean β.

- 88 policyholders submitted 0 claims.
- 12 policyholders submitted 1 claim.
- 3 policyholders submitted 2 claims.
- For two policyholders, it is know that they submitted either 1 claim or 2 claims, but the exact number of claims is not available.
- No policyholder submitted more that 2 claims.

Estimate β using maximum likelihood.

1.5.4 Truncation

For truncated data, the observation is conditional on being outside the truncated range. If data are left truncated at d, such as for a policy with a ordinary deductible of d, so that you only see the observation x if it is greater than d, the likelihood of x is

$$\frac{f(x)}{P(X>d)} = \frac{f(x)}{1-F(d)}$$

For the more rare case of right truncated data — you do not see observations x unless it is under u — the likelihood of x is

$$\frac{f(x)}{P(X<u)} = \frac{f(x)}{F(u)}$$

Example 29.

An auto collision coverage has a deductible of 1200. Claim sizes observed are

$$1300, 1500, 1700, 2300, 3700$$

The data are fitted to an exponential distribution using maximum likelihood. Determine the maximum likelihood estimate of θ.

1.5.5 Combination of Censoring and Truncation

Data that are both left truncated and right censored would have likelihood $\frac{S(u)}{S(d)}$. Grouped data that are between d and c_j in the presence of truncation at d has likelihood $\frac{F(c_j)-F(d)}{S(d)}$.

Example 30.

For an insurance coverage, your department only handles claims below 10,000. Among these claims, you observe 6 claims for 800, 1000, 2000, 2500, 4000, and 5000. In addition, there are 4 claims for amounts below 500, who exact amounts are unknown. You fit these claims to an inverse exponential distribution using maximum likelihood. Determine the resulting estimate for the 66^{th} percentile of claim size for all claims.

Example 31.

You are given the followings:

- A portfolio contains two types of policies.
- Type Y policies have no deductible but a policy limit k.
- Type Z policies have a deductible of k but no policy limit.
- A total of 50 losses that are less than k have been recorded on Type Y policies, $y_1, y_2, ..., y_{50}$.
- A total of 75 losses that exceed k have been recorded on Type Y policies.
- Losses that are less than k are not recorded on Type Z policies.
- A total of 75 losses that exceed k have been recorded on Type Z policies $z_1, z_2, ..., z_{75}$.
- The random variable X underlying the losses on both types of policies has the density function $f(x; \theta)$ and the cumulative distribution function $F(x; \theta)$.

What is the likelihood function of θ.

Example 32 (T1Q13).

7 losses have been recorded in thousands of dollars and are grouped as follows:

Interval	[0, 1)	[1, 5)
Number of Losses	3	4

There is no record of the number of losses at or above 5,000. The random variable X underlying the losses, in thousands, has the density function

$$f(x) = \lambda e^{-\lambda x}, x > 0, \lambda > 0.$$

Determine $L(\lambda)$ and MLE of λ.

1.5.6 Maximum Likelihood Estimators -Special Techniques

1. Cases where MME = MLE (This only applies to complete individual data)
   Poisson distribution, Exponential distribution, Gamma distribution with a fixed α, mean of normal distribution. In these cases, the MLE is the sample mean.

2. For a negative binomial distribution, the MLE of $r\beta$ is the sample mean.

3. Lognormal distribution:
   If $X \sim$ Lognormal(μ, σ^2); then
   $$\hat{\mu} = \frac{\sum_{i=1}^{n} lnx_i}{n},$$
   and
   $$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}[ln(x_i) - \hat{\mu}]^2}{n}$$

Example 33.

You are given the following observations of X:

$$4.9, 1.8, 3.4, 6.9, 4.0.$$

You fit a lognormal distribution to these observations using maximum likelihood method. Determine $\hat{\mu}$ and $\hat{\sigma}$. $\boxed{1.3437,\ 0.1987}$

4. Exponential distribution $X \sim \text{EXP}(\theta)$

$$\boxed{\hat{\theta} = \frac{\sum_{i=1}^{n_{c'}+n_c}(x_i - d)}{n_{c'}}}$$

where $n_{c'}$ is the number of uncensored observations
n_c is the number of censored observations.
x_i are both actual and censored observations,
d is the truncation point.

Example 34.

The random variable X has exponential distribution with mean θ. A random sample of three observations of X yields the values 0.3, 0.55, 0.8. Determine the value of $\hat{\theta}$, the maximum likelihood estimator of θ. $\boxed{0.55}$

Example 35.

A policy has an ordinary deductible of 100 and a policy limit of 1000. You observed the following 10 payments:

$$15, 50, 100, 215, 400, 620, 750, 900, 900, 900.$$

An exponential distribution is fitted to the ground-up distribution function, using maximum likelihood. Determine the estimated parameter θ. $\boxed{692.857}$

Example 36.

For an insurance coverage with claims limit 5, you observe 30 claims for amounts below 5 with sum 45, and 12 claims for 5. You fit this data to an exponential distribution with integral θ using maximum likelihood. Determine θ.

5. Weibull distribution $X \sim \text{WEI}(\tau, \theta)$

$$\hat{\theta} = \left(\frac{\displaystyle\sum_{i=1}^{n_{c'}+n_c} x_i^{\tau} - (n_{c'} + n_c)d^{\tau}}{n_{c'}} \right)^{1/\tau}$$

where
$n_{c'}$ is the number of uncensored observations.
n_c is the number of censored observations.
$x_i's$ are both actual and censored observations,
d is the truncation point.

Example 37 (T1Q14).

For an insurance policy, you are given:

- Ground-up losses follow a Weibull distribution with parameters τ =6 and θ(unknown).
- Losses under 860 are not reported to the insurer.
- For each loss over 860, there is a deductible of 860 and a policy limit of 2000.
- A random sample of six claim payments for this policy is:

$$375 \quad 450 \quad 845 \quad 1080 \quad 1140+ \quad 1140+$$

where + indicates that the original loss exceeds 2000. Determine the 76^{th} percentile of the ground-up distribution.

6. Uniform distribution $(0, \theta)$

If data are individual:

$$\boxed{\hat{\theta} = max(x_1, ..., x_n) = x_{(n)}}$$

If data are censored, truncated or group:

$$\boxed{\hat{\theta} = \frac{(n_{c'} + n_c)u - n_c d}{n_{c'}}},$$

where
$n_{c'}$ is the number of uncensored observations
n_c is the number of censored observations
u is the censored value
d is the truncation point.

Example 38.

You are given the following observations of X:

$$4.9 \quad 1.8 \quad 3.4 \quad 6.9 \quad 4.0$$

You fit a uniform distribution on $[0, \theta]$ to this data using the maximum likelihood method. Determine the MLE of θ. $\boxed{6.9}$

Example 39.

You are given:

- At time 4 hours, there are 5 working light bulbs.
- The 5 bulbs are observed for 15 more hours.
- Three light bulbs burn out at times 5, 9, and 13 hours, while the remaining light bulbs are still working at time 19 hours.
- The distribution of failure times is uniform on $(0, w)$.

Determine the maximum likelihood estimate of w. 29

7. Pareto distribution $X \sim \text{Pareto}(\alpha, \theta(\text{known}))$

$$\boxed{\hat{\alpha} = \frac{-n_{c'}}{K}},$$

where

$$\boxed{K = ln\frac{(\theta+d)^{n_{c'}+n_c}}{\prod_{i=1}^{n_{c'}+n_c}(\theta+x_i)}}$$

for two parameter Pareto distribution

$$\boxed{K = ln\frac{(d)^{n_{c'}+n_c}}{\prod_{i=1}^{n_{c'}+n_c}(x_i)}}$$

for single parameter Pareto distribution and
d is replace by θ if $d < \theta$.

where
$x_i = u$ for $i > n$,
n_c is the number of censored observations.
$n_{c'}$ is the number of uncensored observations

Example 40.
Losses follow a distribution with density function

$$f(x) = \frac{\alpha}{x^{\alpha+1}}, x > 1, \alpha > 0.$$

A random sample of size five produces three losses with values 3, 6 and 14, and two losses exceeding 25. Determine the MLE of α. $\boxed{0.2507}$

1.5.7 Approximate Mean and Variance-Delta Method

Suppose $g(x)$ has derivatives $g'(x), g''(x), \ldots$ in an open interval containing $\mu = E(X)$. The function $g(x)$ has a Taylor approximation about μ,

$$g(x) \approx g(\mu) + g'(\mu)(x-\mu) + \frac{1}{2}g''(\mu)(x-\mu)^2$$

which suggests the approximation

$$\boxed{E[g(x)] \approx g(\mu) + \tfrac{1}{2}g''(\mu)\sigma^2}$$

and using the first two terms,

$$\boxed{V[g(x)] \approx [g'(\mu)]^2\sigma^2}$$

where $\sigma^2 = V(X)$

Suppose $\mathbf{X} = (X_1, X_2, ...X_k)$ is a k dimensional random variable, $\boldsymbol{\theta}$ is the its mean, and $\boldsymbol{\Sigma}$ is its variance covariance matrix. If $g(\mathbf{X})$ is a function of $\mathbf{X}$, then the delta method approximation of the variance is

$$\boxed{V[g(\mathbf{X})] \approx (\boldsymbol{\partial g})'\boldsymbol{\Sigma}(\boldsymbol{\partial g})}$$

where $\boldsymbol{\partial g} = (\frac{\partial g}{\partial x_1}, ..., \frac{\partial g}{\partial x_k})'$ is the partial derivatives of g evaluated at θ, and prime indicates the transpose.

In the one variable case it reduces to

$$\boxed{V[g(X)] \approx (\partial g/\partial X)^2 V(X)}$$

Example 41.
Claim size X follows an exponential distribution with mean θ, $\hat{\theta}$, an estimator for θ, is 1000. This estimator has variance 10,000. Estimate the variance of $P(X < 500)$ when calculate using $\hat{\theta}$. $\boxed{0.000919699}$

Example 42 (T1Q15).

You are given:

- Fifty claims have been observed from a lognormal distribution with unknown parameters μ and σ.
- The maximum likelihood estimates are $\hat{\mu} = 7.2$ and $\hat{\sigma} = 1.64$.
- The covariance matrix of $\hat{\mu}$ and $\hat{\sigma}$ is

$$\begin{bmatrix} 0.0301 & 0 \\ 0 & 0.0106 \end{bmatrix}.$$

Determine the variance of the probability that the next claim will be less than or equal to 4,831 using delta method.

1.5.8 Estimator Quality

Three measures of estimator quality are:

1. **Bias**: $\boxed{B_{\hat{\theta}}(\theta) = E[\hat{\theta}] - \theta}$. An estimator is unbiased if $B_{\hat{\theta}}(\theta) = 0$, which means that based on our assumptions, the average value of the estimator will be the true value, obviously a desirable quality.

   Definition 6.

   A sequence of estimators $\{Tn\}$ is said to be **asymptotically unbiased** for $\tau(\theta)$ if

   $$\lim_{n\to\infty} E(T(n)) = \tau(\theta), \forall\ \theta \in \Omega$$

2. **Consistent**

   Definition 7.

A sequence of estimators $\{Tn\}$ is said to be **consistent** estimators of $\tau(\theta)$ if

$$\forall \epsilon > 0, P[|Tn - \tau(\theta)| < \epsilon] = 1 \ \forall \ \theta \in \Omega.$$

3. **Mean Square error**

$$\boxed{MSE_{\hat{\theta}}(\theta) = E[\hat{\theta} - \theta]^2.}$$

   The lower the MSE, the better the estimator.
   Note: $\boxed{MSE_{\hat{\theta}}(\theta) = V(\hat{\theta}) + B_{\hat{\theta}}(\theta)^2.}$

   Definition 8.

A sequence of estimators $\{Tn\}$ is said to be **mean squared error consistent (MSE consistent)** estimators of $\tau(\theta)$ if

$$\lim_{n\to\infty} E(T(n) - \tau(\theta))^2 = 0 \ \forall \ \theta \ \in \Omega$$

Remarks:

1. If $\{Tn\}$ is MSE consistent, then at also consistent.
2. $\{Tn\}$ is MSE consistent iff it is asymptotically unbiased and $\lim_{n\to\infty} V(T_n) = 0$.

Example 43.
Let $X_1, X_2, \ldots, X_n$ be a random sample with $E(X_i) = \mu$ and $V(X_i) = \sigma^2$. Show that

$$S'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

is a biased estimator for σ^2 and that

$$S'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

is an unbiased estimator for σ^2.

Example 44 (T1Q16)**.**
Two different estimators, ψ and ϕ, are available for estimating the paramentes, β, of a given loss distribution. To test their performance, you have 89 simulated trials of each estimator, using $\beta = 2$, with the following results:

$$\sum_{i=1}^{89}\psi_i = 177, \sum_{i=1}^{89}\psi_i^2 = 392, \sum_{i=1}^{89}\phi_i = 154, \sum_{i=1}^{89}\phi_i^2 = 322$$

Calculate $\frac{MSE_{\psi}(\beta)}{MSE_{\phi}(\beta)}$.

1.5.9 Asymptotic Properties of MLEs

Under certain circumstances, it can be shown that the Maximum Likelihood Estimators are asymptotically unbiased, consistent and are normally distributed, as long as certain regularity conditions hold. i.e. the solutions, $\hat{\boldsymbol{\theta}}$, of maximum likelihood equations have the following properties:

1. $\hat{\boldsymbol{\theta}}$ exists and is unique,
2. $\hat{\boldsymbol{\theta}}$ is consistent estimators of $\boldsymbol{\theta}$.
3. $\hat{\boldsymbol{\theta}}$ is asymptotically normal with asymptotically mean $\boldsymbol{\theta}$ and variance $[I(\boldsymbol{\theta})]^{-1}$

where

For two parameters case:

$$I(\boldsymbol{\theta}) = -nE\begin{pmatrix} \frac{\partial^2 lnf(x;\theta_1,\theta_2)}{\partial\theta_1^2} & \frac{\partial^2 lnf(x;\theta_1,\theta_2)}{\partial\theta_1\partial\theta_2} \\ \frac{\partial^2 lnf(x;\theta_1,\theta_2)}{\partial\theta_1\partial\theta_2} & \frac{\partial^2 lnf(x;\theta_1,\theta_2)}{\partial\theta_2^2} \end{pmatrix}$$
$$= -E\begin{pmatrix} \frac{\partial^2 lnL(x;\theta_1,\theta_2)}{\partial\theta_1^2} & \frac{\partial^2 lnL(x;\theta_1,\theta_2)}{\partial\theta_1\partial\theta_2} \\ \frac{\partial^2 lnL(x;\theta_1,\theta_2)}{\partial\theta_1\partial\theta_2} & \frac{\partial^2 lnL(x;\theta_1,\theta_2)}{\partial\theta_2^2} \end{pmatrix}$$

and for one parameter case:

$$I(\theta) = -nE\left[\frac{\partial^2}{\partial\theta^2} lnf(x;\theta)\right]$$
$$= -E\left[\frac{\partial^2}{\partial\theta^2} lnL(x;\theta)\right]$$

Example 45.
You are given 20 observations from an exponential distribution with mean θ. Calculate the information matrix for this data if $\theta = 5$.

Example 46.
Consider a random sample from a Pareto distribution, $X_i \sim$ Pareto$(\alpha, 1)$. Find the asymptotic distribution of $\hat{\alpha}$, the MLE of α.

Example 47.
A sample of 10 observations comes from a parametric family $f(x,y,\theta_1,\theta_2)$ with log-likelihood function

$$
\begin{aligned}
lnL(\theta_1,\theta_2) &= \sum_{i=1}^{10} lnf(x,y,\theta_1,\theta_2) \\
&= -2.5\theta_1^2 - 3\theta_1\theta_2 - \theta_2^2 + 5\theta_1 + 2\theta_2 + k
\end{aligned}
$$

where k is a constant. Determine the estimated covariance matrix of the maximum likelihood estimator, $\begin{pmatrix}\hat{\theta}_1 \\ \hat{\theta}_2\end{pmatrix}$. □